Ontology-based Testing System for Evaluation of Student's Knowledge

Dauren Kabenov[1], Raihan Muratkhan[1], Dina Satybaldina[1], Bibigul Razahova[1]

[1] L. Gumilyov Eurasian National University

5 Munaitpasov str., 010000 Astana, Kazakhstan

kabenov73@mail.ru, rai-81@mail.ru, dinsaty@gmail.com, utalina@mail.ru

Abstract

In the education system, the control of knowledge is always of primary importance and is complex interaction process between a teacher and a student. Today computer-aided testing of knowledge becomes very popular because the test's results do not depend on the subjective opinion of a teacher. Using traditional tests suggests choosing questions of closed form when several variants of the answer are suggested, one of which is correct and should be selected. It is supposed to use them adequately only in terms of strictly formally asked questions. It leads to absolutely simple questions. But knowledge appropriation includes not only (and not so much) memorizing priori veritable facts but the capability of understanding general phenomena, tendencies. To control this knowledge "open" (without the options) test tasks are more effective. In this paper we have presented an ontology-based text mining approach for the automatic evaluation of a student's answers on the native Language (the Kazakh language). The testing system combined the features of different tests types and formats. The use of intelligent algorithms can also dynamically change the system of evaluation and the test circuit, which significantly improves the quality and speed of testing.

Keywords

Intellectual System; Ontology; Text Mining; Control of Knowledge; Testing Systems

Introduction

Education is one of the priorities of the strategic development program of the Republic of Kazakhstan and has the following major trends: the emergence of the national model of education, the integration of the Kazakhstan education system into the world educational space. The main goal of informatization of the education system is to create a unified educational information environment and to improve the quality of education, ensure equal opportunities for citizens to receive education at all levels and stages, as well as integrate the information space of the Republic of Kazakhstan into the world educational space (Ministry of Education and Science of the Kazakhstan Republic, 2010).

To achieve these goals several problems are to be solved, for example:

- creation of modern e-learning materials in the Kazakh language and the development of support and maintenance;
- development of the modern education methodology on the basis of computer sciences.

Successful e-learning takes place in an integrated system that combines environment tools for creation of electronic educational materials, subsystems for control of knowledge and subsystems for support of the training process. Also practicing credit-module system of the learning process organization makes working out effective means of students' knowledge control important. The research analysis of this problem shows the tendency of enlarging tests usage as an instrument of the studied material quality evaluation (Prisyazhnyuk, 2008). The main advantage of computer tests is the opportunity to ask all the students within the assignment in equal conditions and according to the equal grades scale. It increases the objectivity of knowledge control in comparison with the traditional methods.

Currently, there are many testing systems in various fields of knowledge, for example OLAT (Open Source LMS, 2011), Moodle (Moodle Trust, Moodle.org), Sakai (Sakai Foundation, Sakai Project) and AuthorWare (Adobe Systems Incorporated,). Most of these tools provide the ability to create multimedia tests, testing for traditional learning and e-learning, saving and transfer of results to the teacher for administration of users and educational groups.

On the other hand, using traditional tests suggests choosing answers (and their evaluation correspondingly) on the basis of two-positional logic (Prisyazhnyuk, 2008). It leads to absolutely simple question formulations which “lie on the surface”. To control knowledge “open” questions (without suggested variants of answers) are more effective. However, the adequate automatic verification of answers to the questions of this type is a difficult task. Patterns of the answers in the form of regular expressions are not able to take into account the diversity inherent in the native language. Also the automatic detection of random errors (misprints, typographical errors), and spelling is required.

In order to alleviate this disadvantage of test control of knowledge in this paper, we suggest an application of artificial intelligence methods and tools, in particular, ontological engineering. Description of domain knowledge Planimetry ontology is presented in the paper. Classes of concepts, their structure and properties are defined. List of used relations and characteristics of output procedures are considered.

For texts of the geometric tasks solution ontological descriptions of the situations presented as a result of the transformation and evaluation of the concept's structures. It is shown that the use of ontology-based text mining can open the "anatomy" of the correct answer preparation that can be used in the analysis of a student's answer and the search for precisely that moment, which caused difficulties in their reasoning.

Results of research can be used at creation of intellectual testing systems on the base processing of the Kazakh language. The proposed concept testing system enables the use of intellectual evaluation results of the user level and provides a set of tests, tailored to the level of preparation of the test. The system distributes the issues in terms of complexity, based on data obtained during testing. This enables the construction of adaptive tests which are self-correcting to the level of users.

The paper is organized as follows. Section 2 outlines related work. In section 3 we carry out Planimetry ontology. Section 4 defines the proposed concept testing system based on ontology-based text mining. We finish this paper with some conclusions and future work in section 5.

Related Work

The Computer-aided Testing of Knowledge

For many countries, e-learning is valued and utilized as a driving force to speed up the technical, industrial and economic development of the society. As a research subject, e-learning is both multidisciplinary and interdisciplinary and covers a wide range of research topics, with scholars from different disciplines conducting e-learning related research ranging from content design to associated policy. Longitudinal trends of e-learning research using text mining techniques are described in (Hung Jui-long, 2012). The authors analyze a significant number of research works and provide useful insights into that e-learning research is at the early majority stage and focus has shifted from issues of the effectiveness of e-learning to teaching and learning practices.

Current e-learning theories stress the importance of situated cognition and personalized learning (Hung Jui-long, 2012). Control of knowledge is an intellectual problem, demanding a high-quality solution that will help to reach a new stage in the methodology of teaching, since it could give the opportunity to realize the idea of individual approach to training on a massive scale. The computer-aided testing of knowledge becomes very popular nowadays, firstly, because it saves the working time for teacher, relieves him from routine work and allows providing impartial evaluation of knowledge, the results of which do not depend on the subjective opinion of different teachers.

In (Park, Jooyong., 2010) the newly developed computerized constructive multiple-choice testing system is introduced. The system combines short answer (SA) and multiple-choice (MC) formats by asking examinees to respond to the same question twice, first in the SA format, and then in the MC format.

The authors (Alexeyev A.N., 2007) have developed the software tool that allows preparing test questions and conducting testing using any of the suggested types of questions below. Description of this software tool and the intellectual algorithms for evaluation of knowledge is presented in the previous paper of the authors (Alexeyev A.N., 2003).

The papers (Shahbazova Sh., 2000, Shahbazova Sh.,

2011) describe methods of implementation of a control mechanism of student knowledge with the help of fuzzy set theory combined with neural network technology. The papers apply some serious improvements in the logic of evaluation of knowledge, and methodologies of data interpretation of student responses. The presented architecture is typical of the configuration of hardware and software in an intranet environment of educational institutions.

Analysis of the aforementioned work shows that ordinary linear tests with simple forms of the answer do not quite meet the requirements of comprehensive control of students' knowledge. Most of all concern natural and mathematical sciences, a feature which is the close relationship of concepts, themes and sections of the course, as the main criterion for learning - the ability to solve tasks of different nature and level of complexity. Therefore development of the adaptive, nonlinear, and intellectual testing methods with more different types of tasks and answers' forms are needed. At the same time, new testing systems should incorporate all the achievements of previous generations of the knowledge control tools.

Ontologies

In order to build ontology of Planimetry, it is beneficial to understand the need of ontology and some works connected with the ontology-based text mining.

An ontology is an explicit formal specification of the terms in explicit specifications the domain and relations among them (Gruber T.R., 1995). Ontologies are useful as means to support sharing and reutilization of knowledge (Decker S., 1999). This reusability approach is based on the assumption that if a modeling scheme, i.e., ontology, is explicitly specified and mutually agreed upon by the parties involved, and then it is possible to share, reutilize and extend knowledge. Many disciplines nowadays develop standardized ontologies that domain experts can use to share and annotate information in their fields. Problem-solving methods, domain-independent applications, and software agents use ontologies and knowledge bases built on ontologies as data (Noy Natalya F, 2001).

Reusing existing ontologies may be a requirement if our system needs to interact with other applications that have already committed to particular ontologies or controlled vocabularies (Noy Natalya F, 2001). There are libraries of reusable ontologies on the Web and in the literature, for example, the Ontolingua ontology library (Maedche A., 2003) or the DAML ontology library (Available from:http://www.daml.org /ontologies).

The need of ontologies is connected with the inability of the existing methods to adequately automatically process native-language texts. For high-quality word processing, you must have a detailed description of the problem area with a lot of logical links that show the relationships between the terms field. The use of ontologies can provide a native language text in such a way that when it becomes available-for-automatic processing (Available from: http://www.ksl.stanford. edu/software/ ontolingua/).

In this paper we develop the Planimetry ontology and this ontology is used as a basis for the automatic verification of geometric task solution. The authors (Khakhalin G., 2012) have developed the similar ontology for automatic synthesis of structural images of the planimetric figures. We use some concepts of the ontology, which get the authors consent.

Planimetry Ontology

An ontology is a formal explicit description of concepts in a domain of discourse (*classes* (sometimes called *concepts*)), properties of each concept describing various features and attributes of the concept (*slots* (sometimes called *roles* or *properties*)), and restrictions on slots (facets (sometimes called *role restrictions*)) (Noy Natalya F, 2001). Ontology together with a set of individual instances of classes constitutes a knowledge base.

Development of an ontology includes (Noy Natalya F, 2001):

- defining classes in the ontology,
- arranging the classes in a taxonomic (subclass–superclass) hierarchy,
- defining slots and describing allowed values for these slots,
- filling in the values for slots for instances.

Our ontology is structured in three levels. The first level contains classes whose instances cannot be derived from *other* classes. Class Plane Shape represents all objects of planimetric shapes. Specific Plane Shapes - instances of this class.

The next levels are generated through a reasoning process, that is, using the ontology reasoner or through the different modules called by the Reasoning Manager. The higher the level is; the more detailed the information is, e.g., in the second level polygons are classified into pentagons or triangles or quadrangles and in the third level the triangle are further classified into rectangular, equilateral and isosceles triangle (if the classification by sides will be considered, see Fig. 1).

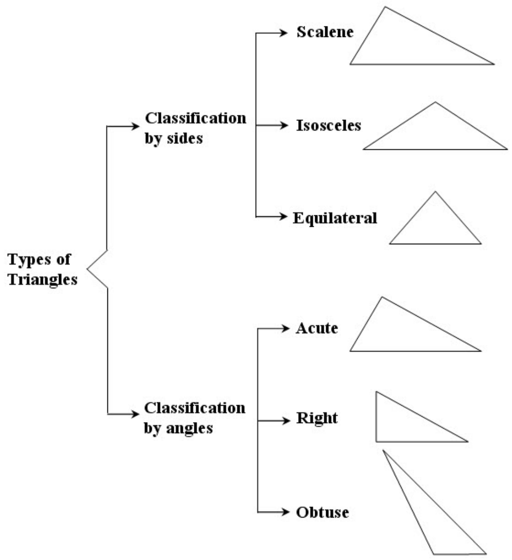

FIG. 1 TYPES OF TRIANGLES. CLASSIFICATION BY SIDES AND CLASSIFICATION BY ANGLES.

We organize the classes into a hierarchical taxonomy by asking if by being an instance of one class, the object will necessarily (i.e., by definition) be an instance of some other class.

If a class A is a superclass of class B, then every instance of B is also an instance of A.

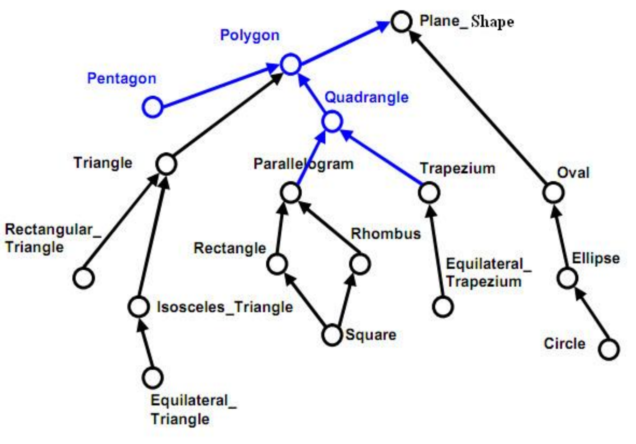

FIG. 2. SOME CLASSES AND INSTANCES IN THE PLANE_SHAPE DOMAIN. DIRECT LINKS REPRESENT INSTANCE-OF AND SUBCLASS-OF.

In other words, the class *B* represents a concept that is a "kind of" *A* (Noy Natalya F., 2001). For example, every equilateral triangle is necessarily an isosceles triangle. Every isosceles triangle is necessarily a triangle. Therefore the equilateral triangle class is a subclass of the equilateral triangle class. A fragment of the taxonomy for the Planimetry ontology is presented in Fig. 2.

Slots describe properties of classes and instances. Each property can be set to a specific value or a formula to calculate this value of the property: for example, a triangle has sides, angles, sides with lengths, angles with a degree measure. All subclasses of a class inherit the slot of that class. On the other hand, subclasses can have their own properties. For example, an isosceles triangle has sides that have a length, with those two sides (legs) being congruent, and the third party has a special name - base (see Fig.3 and Fig. 4).

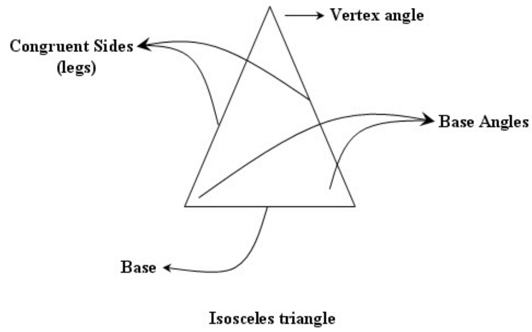

FIG. 3. ISOSCELES TRIANGLES. THE ANGLES OPPOSITE THE TWO CONGRUENT SIDES ARE CALLED BASE ANGLES AND THE BASE ANGLES. THE ANGLES OPPOSITE THE BASE IS CALLED VERTEX ANGLE.

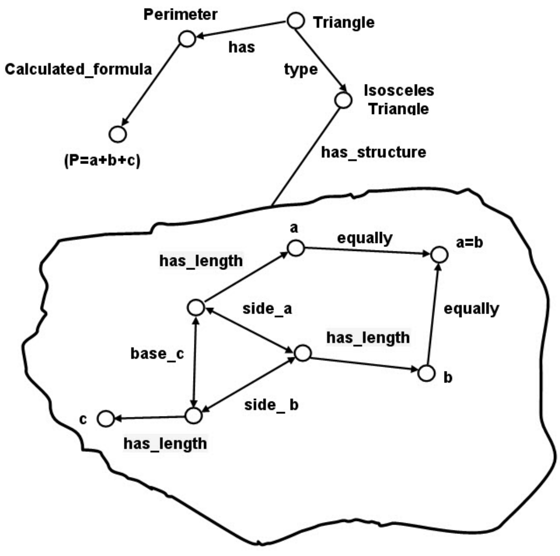

FIG. 4. STRUCTURE OF THE AN ISOSCELES TRIANGLE CONCEPT TO PARSE OF THE SOLUTION OF THE GEOMETRIC TASKS.

Fragments of the ontology, including the structure and properties of concepts are the basis for description of the situation, which is determined by the input data to solve a geometrical task. The concepts and relationships defined by the input conditions (the geometrical task's text) are introduced in addition to these ontology fragments.

Proposed Concept Testing System Based on Ontology-based Text Mining

Compared to traditional forms of learning, e-learning has several advantages: adaptive to the individual characteristics of students, the freedom to choose the time, place and level of education, use of new teaching methods, modern means of communication and information transfer between a student and a teacher. However, control of knowledge is particularly important because of the lack of direct contact between student and teacher.

Organization of control of knowledge is closely connected with the problem of selecting the type of questions, mode of the testing trajectory formation and methods of the answers verification.

To solve these problems we propose the concept of intellectual testing system on the base of the domain knowledge ontology. Planimetry ontology is used as the domain knowledge ontology.

The following types of questions are offered for the control of knowledge quality:

- test questions of closed form, i.e., when several variants of the answer are suggested, one of which is correct and should be selected;
- test questions of open form, i.e., questions without suggested variants of answers (such questions are useful for evaluation of knowledge of terms, definitions, notions, etc);
- situation tests, i.e., set of test assignments designed for solution of problematic situations (a geometrical task).

A special method of selection of a testing trajectory mode is proposed. Test set is not formed by random sample from tests database. Questions selection is based on the analysis of answers to previous questions. This algorithm is based on the original method of choosing questions according to the system, corresponding to the current level of student knowledge (Shahbazova, Sh., 2011).

For analysis and verification of student's answers to test questions of open form ontology-based text mining is proposed. Description of domain knowledge ontology of Planimetry was presented in Section 3. Semantic analysis of the native language texts is the following stage.

1) Pre-linguistic processing of the source text (morphological and syntactic analysis of sentences) is needed to separate terms (classes, subclasses, properties and relations).

2) A formal understanding of the text as result of constructing an ontological graph.

In paper (Razakhova B., 2009) the formalized syntactic rules, analysis and synthesis algorithms of word-combinations and sentences of the Kazakh language were constructed. The results of research can be used at creation of intellectual human-machine systems with interaction possibility in the Kazakh language.

Let us describe in brief the proposed method of verification of the geometric task solution on the basis of ontology. Texts of geometrical task are a set of connected sentences (Khakhalin G., 2012). These include simple and complex sentences, incomplete sentences (with an anaphora and an ellipsis). A formal understanding of the text of geometrical problems is their representation in the language of domain knowledge ontology of Planimetry. This representation must be connected and extended with filling in the values for slots for instances from the description of the situation presented by the text.

Considering the whole process of analysis with geometric problems in this paper is not possible. Therefore, let us consider the situation's structure which should be obtained as a result of ontology-based linguistic analysis for one geometric task.

Text of task: determine the type of triangle that has sides of length which is equal to 5, 6, 6.

Pre-linguistic processing of the source text will give the following concepts and combinations of concepts that are comparable to the ontology presented in Figure 4, for example: type (of triangle); having sides (from triangle class); and length, is equal.

For the formal description of Planimetry ontology Prolog (a logic programming language) is used (Visual Prolog 7.3). Prolog's inference engine is used to build an ontological graph. A formal description of the

ontology (see Fig. 4) uses the following data types with alternative functors.

```
domains
poligon = triangle;
      pentagon;
      quadrangle
triangle=triangle1(classification_by_sides);
       triangle2(classification_by_angles);
classification_by_sides=scalene;
                  isosceles_triangle;
 equilateral_triangle scalene=sc(sides)
isosceles_triangle=is_tr(sides)
equilateral_triangle=eq_tr(sides)
sides=sides(side_a, side_b, side_c)
side_a, side_b, side_c =real
classification_by_angles=acute_triangle;
                   right_triangle;
                   obtuse_triangle
acute_triangle=ac_tr(angles)
right_triangle=r_tr(angles)
obtuse_triangle=o_tr(angles)
angles=angles(angleA, angleB, angleC)
angleA, angleB, angleC =real
predicates
nondeterm treangle_tip(treangle,symbol)
```

Hence, an inference rule describing that the triangle is isosceles can be given by

```
treangle_tip(treangle(sides(Side_a,Side_b,Side_c),
   angles(Angle_A,Angle_B,Angle_C)),Y):-Side_a=Side_b,
Angle_A=Angle_B,Y=isosceles;
Side_b=Side_c,Angle_B=Angle_C,Y=isosceles;
   Side_a=Side_c,Angle_A=Angle_C,Y=isosceles,!.
treangle_tip(treangle(sides(Side_a,Side_b,Side_c),
   angles(Angle_A,Angle_B,Angle_C)),Y):-
   Side_a<>Side_b,Side_b<>Side_c,
   Angle_A+Angle_B+Angle_C=3.14,Y=scalene,!.
treangle_tip(treangle(sides(Side_a,Side_b,Side_c),
   angles(Angle_A,Angle_B,Angle_C)),Y):-
   Side_a=Side_b, Side_b=Side_c,Y=equilateral;
   Angle_A=3.14/3,Angle_B=3.14/3,
   Angle_C=3.14/3, Y=equilateral.
treangle_tip(treangle(sides(Side_a,Side_b,Side_c),
   angles(Angle_A,Angle_B,Angle_C)),Y):-
   Angle_A<1.57,Angle_B<1.57, Angle_C<1.57,
   Angle_A+Angle_B+Angle_C=3.14,Y=acute.
treangle_tip(treangle(sides(Side_a,Side_b,Side_c),
    angles(Angle_A,Angle_B,Angle_C)),Y):-
    Angle_A=1.57, Angle_A+Angle_B+Angle_C=3.14,
    Y=right; Angle_B=1.57,
    Angle_A+Angle_B+Angle_C=3.14,
    Y=right;Angle_C=1.57,
    Angle_A+Angle_B+Angle_C=3.14,Y=right.
treangle_tip(treangle(sides(Side_a,Side_b,Side_c),
    angles(Angle_A,Angle_B,Angle_C)),Y):-
   Angle_A>1.57,Angle_A+Angle_B+Angle_C=3.14,
   Y=obtuse; Angle_B>1.57,
   Angle_A+Angle_B+Angle_C=3.14,Y=obtuse;
   Angle_C>1.57,Angle_A+Angle_B+Angle_C=3.14,
   Y=obtuse.
```

Conclusions and Future Work

This paper presents an ontology-based approach that addresses the problem of ontology-based text mining. We have developed Planimetry ontology for automatic verification of answers to the questions of "open" questions (without options). Results of the

research can be used in creation of intellectual testing systems on the base of the Kazakh language processing.

Our conception of the testing system uses intellectual evaluation results of level of the user and provides a set of tests, tailored to the level on which the student prepared. The control knowledge system distributes the issues in terms of complexity, based on the data obtained during testing. This enables the construction of adaptive tests which are self-correcting to the level of users.

Future research in the frame of methodological aspects of computer-aided control knowledge will be concerned to development of the test questions databases of different types and different levels of complexity. Also in the technical aspects frame researches will continue development software that implements of all the stages of semantic analysis of texts on the basis of the ontological engineering and native languages processing.

REFERENCES

Alexeyev A.N., Volkov N.I., Kochevsky A.N., Elements of Fuzzy Logic in Computer-Aided Control of Knowledge, Otkrytoye obrazovaniye, Vol. 4, Moscow, 2003; pp.23–25 (In Russian).

Alexeyev, A.N., Volkov, N.I., Kochevsky, A.N., Types of Questions for Computer-Aided Testing of Students' Knowledge Computer Modelling and New Technologies, Vol.11.3, Riga, 2007; pp.35–42.

Adobe Systems Incorporated, Adobe Authorware 7, [Online] Available from:http:// www.adobe.com /products /authorware/

Decker S., Erdmann M., Fensel D. and Studer D., Ontobroker: Ontology based access to distributed and semi-structured information, DS-8: Semantic Issues in Multimedia Systems, 1999.

Gruber T.R., Toward principles for the design of ontologies used for knowledge sharing, International. J. of Human-Computer Studies, Vol.43(5-6), 1995; pp.907–928.

Hung Jui-long., Trends of E-learning Research from 2000 to 2008, Use of text mining and bibliometrics, British J. of Educational Technology, Vol. 43.1, 2012; pp.5–16.

Khakhalin G., Kurbatov S., Naidenova K., Lobzin A., Integration of the Image and NL-text Analysis/Synthesis Systems. In Rafael Magdalena (eds): Intelligent Data Analysis for Real-Life Applications: Theory and Practice, IGI Global, 2012.

Maedche A., Neumann G. and Staab S., Bootstrapping an Ontology-based Information Extraction System, Intelligent exploration of the web, Physica-Verlag GmbH Heidelberg, Germany, 2003; pp.345 – 359.

Ministry of Education and Science of the Kazakhstan Republic., State Program of Education Development in the Republic of Kazakhstan for 2011–2020. Approved by the Decree of the President of the Kazakhstan Republic, Astana, 2010, № 1118.

Moodle Trust, Moodle.org, Open-source Community-based Tools for Learning. [Online] Available from: http://moodle.org/.

Noy Natalya F. and McGuinness, Deborah L., Ontology Development 101, A Guide to Creating Your First Ontology', Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

OLAT – The Open Source LMS, [Online] Available from: http://www.olat.org/ website/en/ html/index.html.

Park, Jooyong., Constructive multiple-choice testing system, British J. of Educational Technology, Special Issue, Learning objects in progress, Vol. 41.6, 2010; pp.1054–1064.

Prisyazhnyuk, E., Fuzzy model of the Automated System of Module Knowledge Control, Int.l J. Information Technologies and Knowledge, Vol.2, 2008; pp.465-468.

Razakhova B., Automation of the analysis and synthesis of word-combinations and sentences of the Kazakh language, PhD thesis (in Kazakh language), L. Gumilyov Eurasian National University, Kazakhstan, 2009.

Sakai Foundation, Sakai Project, [Online] Available from: http://sakaiproject.org/.

Shahbazova Sh., Freisleben B.A., Network-Based Intellectual Information System for Learning and Testing, Proceedings of the Fourth International Conference on Application of Fuzzy Systems and Soft Computing, Siegen, Germany, 2000; pp.308-313.

Shahbazova, Sh.: Application of Fuzzy Sets for Control of Students Knowledge. Application Comput. Math., Vol.10.1, Special Issue, 2011; pp.195-208.

The DAML ontology library, [Online] Available from: http://www.daml.org/ontologies

The Ontolingua ontology library, [Online] Available from: http://www.ksl.stanford.edu/software/ ontolingua/

Visual Prolog 7.3. [Online] Available from: http://www.visual-prolog.com.

Dauren Kabenov in 1997 graduated from the Pavlodar state university of S.Toraigyrov (Kazakhstan) on a speciality "Mathematics and computer ingenering". In 2007 graduated from a magistracy on a speciality "Computer science". He is currently a PhD doctoral chair, "Theoretical Computer Science" of the Eurasian National University. L.Gumilev (Kazakhstan). Area of interests: artificial intelligence, the theory of indistinct sets and fuzzy logic.

Dina Satybaldina got a Candidate of Physics and Mathematics Sciences degree (PhD) in 2000 from the Karaganda State University, Kazakhstan. In 2011, she got a PhD degree in Computer Science from the Al-Farabi Kazakh National University, Kazakhstan. Now she is an Associate Professor of Computer Engineering Department at the L.Gumilyov Eurasian National University, Astana, Republic of Kazakhstan. Her research interests include Cryptography, Cryptography Engineering, Artificial Intelligent and Fuzzy Neural Network. Award Grants of Education and Sciences Ministry of the Kazakhstan Republic for young talented scientists (2000-2002).

Raihan Muratkhan in 2002 graduated from E.A.Buketov Karaganda state university (Kazakhstan) on peciality "Applied mathematics". In 2004 he graduated from magistracy - "Applied mathematics". At present time he is a PhD doctoral candidate of "Theoretical Computer Science" chair of the L.Gumilev Euroasian National university (Kazakhstan). Area of interests: information security, risk management of information security, theory of indistinct (fuzzy) sets and indistinct (fuzzy) logic.

Bibigul Razakhova got a Candidate of Technical Sciences degree in 2009 from the L. Gumilyov Eurasian National University, Kazakhstan. Now she is an Associate Professor of Theoretical Informatics Department at the L.Gumilyov Eurasian National University, Kazakhstan. Her research interests include Artificial Intelligence, Computational Linguistics, Processing Natural Language, Application of Intelligent system and Fuzzy Neural Network for assessment knowledge.